DOCUMENT RESUME

ED 151 346 SP 012 359

AUTHOR Coker, Homer
TITLE An Effectiveness-Directed Approach to Teacher Evaluation and Certification.
PUB DATE Feb 78
NOTE 14p.; Paper presented at the Annual Meeting, American Association of Colleges for Teacher Education (Chicago, Illinois, February 20-24, 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Achievement Rating; *Evaluation Methods; *Observation; *Performance Based Teacher Education; Preservice Education; Student Teachers; Teacher Behavior; *Teacher Certification; *Teacher Evaluation; *Teacher Qualifications; Teaching Skills

ABSTRACT

This paper describes a project to develop an alternate approach to the certification of teachers other than the approval of transcripts submitted by individual candidates or of the programs offered by institutions engaged in teacher preparation. The purpose of the project was to identify a set of "generic competencies" and to validate them in terms of student growth. Special task forces of classroom teachers from the school district in which the study was conducted spent a year meeting, discussing, and studying various definitions of teacher competence. Eleven competency statements were chosen, and specific teacher and student behaviors were listed under each competency area. A group of qualified active educational researchers from the area of teacher observation were selected as consultants. Also, three regular teachers from the school district were trained and employed as observers. Observation of classroom behaviors in the selected competencies was conducted for a full school year. The performance of student teachers was evaluated in the context of achieving skills in the stated competencies, and recommendation for certification was based upon these evaluations. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED151346

AN EFFECTIVENESS-DIRECTED APPROACH TO TEACHER EVALUATION AND CERTIFICATION

HOMER COKER

WEST GEORGIA COLLEGE
CARROLLTON, GEORGIA 30118

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Homer Coker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

SP012 359

A paper read at a workshop
American Association of Colleges for Teacher Education
Chicago, Illinois
February 22, 1978

AN EFFECTIVENESS-DIRECTED APPROACH TO TEACHER EVALUATION AND CERTIFICATION

Homer Coker
West Georgia College
Carrollton, Georgia 30117

THE CBTC PROJECT

In 1973, a project was begun in Georgia to develop an alternate approach to the certification of teachers other than the approval of transcripts submitted by individual candidates or of the programs offered by institutions engaged in teacher preparation.

The Georgia State Department of Education selected West Georgia College, a senior college of the University system of Georgia, and the Carroll County Public Schools to help develop and test this alternative approach.

The purpose of the project was to identify a set of "generic competencies" and to validate them in terms of student growth. Since the major task was to develop instrumentation and procedures to be applied to prospective teachers--these would be the basis for licensing--fundamental decisions were made which established certain safeguards.

The first decision was to have the teachers themselves identify broad competency areas. Special task forces of classroom teachers from the school district in which the study was conducted spent a year meeting, discussing, and studying various definitions of teacher competence. They finally chose eleven competency statements. More importantly, they also listed specific teacher and student behaviors under each competency area whose presence

in the classroom should indicate that the teacher possessed each competency. This was important because there is considerable evidence that we should be seeking a large number of small effects rather than a small number of large effects (Soar, 1976) Also, it appears that teacher evaluation should not be based on a global concept of teacher effectiveness, but rather on the idea that an effective teacher is one who possesses a large repertory of competencies (knowledge, skills, etc). There is not necessarily a complete list of competencies such that when a teacher possesses them all, he is fully competent. This is not consistent with the idea that a teacher's effectiveness can improve throughout his career (Medley, 1977)

Our approach to teacher evaluation is based on the identification of a list of competencies of behaviors, performances, etc., which have been shown to correlate with pupil learning. When we talk about evaluating a teacher, we mean we are finding out which of these competencies he possesses and which he does not.

The next decision was to try out only behavioral indicators that had been identified beforehand as promising instead of blindly testing all possible indicators. The chances that a behavior which is found to predict teacher effectiveness will stand up on cross-validation should be greater when only a limited number of promising behaviors are tested.

The competencies listed are not very different from those on other lists that have been developed by other groups as specifications of C/PBTE programs; what is unusual about this

list is that each competency is defined in terms of specific, observable behavioral indicators. It is these indicators that formed the basis for the development of the measures described and that will also be the basis of the performance test yet to be developed.

The third decision was not to attempt to construct an observation instrument specially designed to measure these behavioral indicators, but to select a small number of tested instruments. The reason behind this decision was that building a valid observation system is not as easy as it looks. Too many similar projects have underestimated the time and effort such an enterprise entails; and too many have ended up with a mass of data that are largely uninterpretable. In order to avoid this particular trap, we decided to use existing instruments of proven quality - those which were the product of 15 to 20 years of development and refinement. All had been used in more than one previous study by more than one investigator and they represented several distinct points of view about classroom behavior and about observational methodology. It seemed to us that almost any indicator we might care to study could be identified on one or more of these instruments (and we found this to be the case). (Medley, 1977)

The fourth decision was that the instruments would be low-inference rather than high-inference. The main reason for this was practical: high-inference ratings are more vulnerable to legal attacks in the courts. Sooner or later a candidate who fails on the instrument will go to court and question its validity. If the decision to fail him has been based on high-

inference ratings, the defense must ultimately be based on the expertise of the rater, since his judgments in the form of ratings will constitute the sole documentary evidence that the candidate was incompetent. If the decision is based on low-inference performance observations, however, an objective record of what behaviors the candidate did and did not exhibit will be in the record, together with a competency score derived from that record by the use of the same scoring key that was used on every other candidate. Thus any charge of bias or of unreliability can be refuted, and a strong case built against any charge of invalidity (Medley, 1975).

The other advantages of low-inference instruments are well-known and will not be recapitulated here.

The fifth decision was to select as consultants a group of qualified active educational researchers from the area of teacher observation. This lead us to employ a distinguished group of experts who have been and are currently involved with this effort.

We also selected three(3) regular teachers from the school district, trained them in the several instruments and employed them as observers.

The sixth decision was to assess pupil outcomes and the presence or absence of these teachers behaviors through observation in the natural setting over a full school year. We recruited sixty(60) teachers for the first year of observation and forty-three(43) for the replication study. The teachers attended summer workshops where they were given the set of competencies and the complete instruments.

The final decision was to base each indicator on more than one single item whenever possible. Cronbach (1951) has shown how rapidly a score based on a set of test items with low inter-correlations increases in internal consistency as the number of items increases--a phenomenon familiar to every measurement student as the Spearman-Brown Law.

Although scores based on individual items of behavior are extremely unstable and ambiguous as to meaning, the combination of as few as four or five into a composite can produce a dramatic increase both in stability and in internal consistency, with an accompanying reduction in ambiguity (Medley and Mitzel, 1963, pp. 308-309).

The Instruments

Among the instruments utilized to assess various aspects of teacher performance were the following:

OScAR 5V: Observation Schedule and Record, Form 5 Verbal (Medley). This instrument is designed to analyze and record teacher-student cognitive interactions so that a clear analysis can be made of the classroom's verbal learning environment. OScAR 5V represents more than 20 years work on the OScAR series and has been used extensively in classroom observational research since the 1950's.

STARS: Spaulding Teacher Activity Rating Schedule (Spaulding). This instrument examines 25 categories of teacher behaviors and 19 categories of student behaviors - forming a 19x25 matrix - including teacher listening and reinforcement patterns, concept attainment and concept checking, as well as affective and valuing behaviors.

FLACCS: Florida Climate and Control System (Soar, Soar, and Ragosta). This instrument focuses on classroom direction and control by the teacher and student response. FLACCS records positive and negative affect, non-verbal expressions of affect, classroom physical movement, and the dominance of individual or group activities.

TPOR: Teacher Practices Observation Record (Brown)

This instrument provides a model by which the observer can check off methods of instruction used by the teacher. TPOR is concerned primarily with cognitive learning patterns and focuses on questions such as open-ended versus closed-ended questioning, reinforcing teacher behavior, strong control by the teacher versus a strategy of following class interests, tight versus open structure of conversations, etc.

CASES: Coping Analysis Schedule for Educational Settings (Spaulding). This instrument provides measurements of student personality development and socialization in structured settings. It focuses on the range of student behavior in relation to the teaching strategy. These student behaviors have been empirically refined in case studies of more than 2,000 children over several years.

Other Instruments:

Among the instruments which were utilized to assess student growth were achievement and self-concept measures and a measure of student behavior in the classroom as well as an index of student socioeconomic status.

Data Collection

These instruments were applied by trained observers in 103 classrooms in a single school district.

Records were made of the incidence in each classroom on over 1,300 possible behavioral indicators. In addition, 58 different measures of outcomes were also obtained. This mass of data permits the calculation of over 80,000 process-product correlations. However, by carefully selecting fewer than 100 behavioral indicators from the 1,300 to be correlated with outcomes, we reduced the total number of correlations substantially.

Conclusions

Findings from this project have shown that effective teaching can be predicted from observation in natural classroom settings. The project has demonstrated that a regular classroom teacher released for the purpose of observation can be trained to use a number of observation instruments; that two or more teachers so trained can assess another teacher's performance independently and at different times, with a level of agreement high enough to provide reliable data on that teacher's performance.

There is preliminary evidence that objective scoring keys can be applied to the records of observations in a teacher's classroom tha will yield scores which are valid predictors of how much his pupils are learning from him. There is evidence that scores on such records correlate with pupil gains in reading, for example, about as well as the SAT correlates with college achievement. The important fact to note is that these scores are

objective: a teacher's scores, unlike supervisor's ratings, are in no way dependent on the judgment of the observers. All an observer needs to be able to do is to record which of the events and procedures on a predetermined list occur in a classroom during a certain period of time. He is not required to make any judgments about the appropriateness of the methods used, the quality of the questions asked, the clarity of the explanations, etc.

Computer programs have been devised which will compare the performances recorded in a teacher's classroom with those observed in the classrooms of effective teachers and print out scores which indicate how closely the teacher in question approximates the behavior of an effective teacher (when we identify one).

The teacher effectiveness records will be detailed enough so that any discrepancy between observers will be immediately visible. Evidence indicates that such discrepancies will be minor. When this system is used, a detailed and accurate record will exist of each behavior that affected a teacher's score. It will, therefore, be possible to inform the prospective teacher who is not certifiable just how he needs to change his behavior in order to become certifiable.

There will be no risk that a prospective teacher may draw a rater who is biased, careless or one whose standards may have deteriorated by the time he visits this prospective teacher. Evidence of the reliability of this approach indicates that <u>which</u> observer sees a teacher (or <u>when</u>) has very little effect on the score. If a candidate wishes to appeal his case (or to sue the state) documentary record of his performance would be available

along with the scoring keys, validity data etc. A prospective teacher who failed would receive a diagnostic profile and a prescription for self-improvement and would become eligible for reassessment after a minimum period of time.

Annual reports could be made to the institutions regarding the overall quality of the entire group of prospective teacher examined each year indicating areas of strength and weakness and areas of improvement or decline from year to year. Implementation of this system should substantially improve the quality of teachers coming from the teacher education programs (Medley, Coker, Lorentz, 1976)

The ultimate responsibility of deciding what student teachers should acquire belongs to the teacher educators - those who do the training define the competency by developing an operational definition of the behavior. The instrument must contain the variables which respond to any set of performance competencies (behaviors) which are selected as important and worthy of testing.

We have reduced the items from the five observation schedules into a composite instrument called Georgia Assessment of Teacher Effectiveness (GATE, 1977). We chose to develop a sign instrument because of its greater flexibility, higher objectivity, shorter period required for observer training, and it is more error free.

The initial analysis further suggests that the teaching-learning relationship is much more complex than most educational theorists sometime imply. Each grade level and subject matter requires a different approach. Not only does socioeconomic status affect the student's behavior, there is evidence that it directly

influences the teacher's behavior, as well. Finally, teaching may be like cooking in that some competencies are required in large doses while only a pinch of others may be needed. Just as proportions and preparation times change from recipe to recipe, so also does the required competencies change from setting to setting.

Moreover, the initial data indicate that simple statements of so-called generic competencies will be worse than useless; they will dangerously mislead teachers and school districts into activities that are good in some classes and counterproductive in others.

Travers presented a paper at the 1974 Annual Meeting of AACTE in which he urged the establishment of "Empirically Based Teacher Education". He suggested that the term, effectiveness directed, seem to contain the right substance for a name to describe a program of teacher education based on what is known about effective teaching in the classroom. His recommendation bears repeating. (Travers, 1974)

We have attempted to utilize this approach in our studies, in both preservice/inservice and certification efforts.

Sir Robert Watson-Watt, inventor of radar, is reported to have said "give them the third best to go on. The best never comes, and the second best comes too late". We feel that we have reached a point which approximates the "third best" level.

REFERENCES

Brown, B. B. "Experimentalism in Teaching Practice," Journal of Research and Development in Education, 1970, 4, 14-22.

Coker, H., Lorentz, J. L., & Coker, J. G. Interim Report on Carroll County CBTC Project, Fall 1976. Georgia State Department of Education, 1976.

Comprehensive Tests of Basic Skills. Monterey, California: McGraw-Hill, 1974.

Frymier, J. R. "Development and Validation of a Motivation Index," Theory into Practice, 1970, 9 (1), 56-88.

Gordon, I. J. A Test Manual for the How I See Myself Scale. Gainesville: Florida Educational Research and Development Council, 1968.

Hieronymus, A, N., & Lindquist, E. F. Iowa Tests of Basic Skills. Boston: Houghton Mifflin Company, 1971.

Hieronymus, A. N., Lindquist, E. F., & Others. Iowa Tests of Basic Skills. Boston: Houghton Mifflin Company, 1972.

Lorentz, J. L. The Development of Measures of Teacher Effectiveness from Multiple Measures of Student Growth. A paper presented at the 1977 Annual Meeting, American Educational Research Association, 1977.

Medley, D. M. Teacher Competence and Teacher Effectiveness: A Review of Process-Product Research. PBTE publication. Washington D.C. American Association of Colleges for Teacher Education, 1977.

Medley, D. M. An Approach to the Definition and Measurement of Teacher Competency. A paper presented at the 1977 Annual Meeting, American Educational Research Association, 1977.

Medley, D. M. Observation Schedule and Record, Form 5, Verbal. Mimeographed, 1973.

Medley, D. M. & Mitzel, H. Handbook of Research on Teaching, ed. N.L. Gage, Chicago: Rand McNally, 1963, p.308, 309.

Medley, D. M., Soar, R. S. & Soar, R. M. Assessment and Research in Teacher Education. PBTE Series No.17. Washington, D.C. American Association of Colleges for Teacher Education, 1975.

Monroe, M., Manning, J. C., Wepman, J. M., & Gibb, G. E. The Scott, Foresman Initial Survey Test. Glenview, Illinois: Scott, Foresman and Company, 1972.

Showers, Beverly. The Use of Classroom Observation Measures in the Design of an In-Service Training Program. A paper presented at the Annual Meeting of the American Educational Research Association, 1976.

Soar, R. S. Using Process/Product Relationships as the Basis of Describing Effective Teacher Behavior. A paper presented at the 1977 Annual Meeting. American Educational Research Association, 1977.

Soar, R. S. Final Report--Follow Through Classroom Process Measurement and Pupil Growth, (1970-71). Gainesville: Institute for Development of Human Resources, University of Florida, 1973.

Soar, R. S., Soar, R. M. & Ragosta, M. J. Florida Climate and Control System: Observer's Manual. Gainesville, Fla. Institute for Development of Human Resources, University of Florida, 1971.

Spaulding, R. L. Classroom Behavior Analysis and Treatment, San Jose: Author, 1970.

Spaulding, R. L. A Category System (CASES) and an Interaction Analysis System (STARS/CASES) in Relation to Teacher Competencies. A paper presented at the Annual Meeting, AERA, 1976.

Travers, R. M. W. "Empirically Based Teacher Education", paper presented at AACTE Annual Meeting, 1974.

Warner, W. L., Meeker, M., & Eells, K. Social Class in America, New York: Harper, 1960.